

ABSTRACT OF THE INVENTION

A computer-implemented method for processing a plurality of toponyms, the method involving: in a large corpus, identifying geo-textual correlations among readings of the toponyms within the plurality of toponyms; and for each toponym selected from the plurality of toponyms, using the identified geo-textual correlations to generate a value for a confidence that the selected toponym refers to a corresponding geographic location. Also a method of generating information useful for ranking a document that includes a plurality of toponyms for which there is a corresponding plurality of (toponym,place) pairs, there being associated with each (toponym,place) pair of said plurality of (toponym,place) pairs a corresponding value for a confidence that the toponym of that (toponym,place) pair refers to the place of that (toponym,place) pair. This further method includes, for a selected (toponym,place) pair of the plurality of (toponym,place) pairs, (1) determining if another toponym is present within the document that has an associated place that is geographically related to the place of the selected (toponym, place) pair; and (2) if a toponym is identified within the document that has an associated place that is geographically related to the place of the selected (toponym, place) pair, boosting the value of the confidence for the selected (toponym,place) pair.